Frequency Reuse of Beam Allocation for Multiuser Massive MIMO Systems

Junyuan Wang, *Member, IEEE*, Huiling Zhu, *Member, IEEE*,
Nathan J. Gomes, *Senior Member, IEEE*, and Jiangzhou Wang, *Fellow, IEEE*

***Abstract*—Massive multiple-input-multiple-output (MIMO) has become a promising technique to provide high-data-rate communication in fifth-generation (5G) mobile systems, thanks to its ability to form narrow and high-gain beams. Among various massive MIMO beamforming techniques, the fixed-beam scheme has attracted considerable attention due to its simplicity. In this paper, we focus on a fixed-beam based multiuser massive MIMO system where each user is served by a beam allocated to it. To maximize the sum data rate, a greedy beam allocation algorithm is proposed under the practical condition that the number of radio frequency (RF) chains is smaller than the number of users. Simulation results show that our proposed greedy algorithm achieves nearly optimal sum data rate. As only the sum data rate is optimized, there are some "worst-case" users who could suffer from strong inter-beam interference and thus experience low data rate. To improve the individual data rates of the worst-case users while maintaining the sum data rate, an adaptive frequency reuse scheme is proposed. Simulation results corroborate that our proposed adaptive frequency reuse strategy can greatly improve the worst-case users' data rates and the max-min fairness among served users without sacrificing the sum data rate.**

***Index Terms*—Frequency reuse, beam allocation, achievable data rate, worst-case users, massive multiple-input-multiple-output (MIMO)**

I. Introduction

To support various high-data-rate mobile applications such as online-gaming and high-definition video streaming, massive multiple-input-multiple-output (MIMO) has been proposed as a promising technique for the fifth-generation (5G) mobile communication system [1]. In a massive MIMO system, a large number of antennas are deployed at the base-station (BS), thanks to which a few advantages can be harvested, e.g., the random channel vectors from users to the BS become pairwisely orthogonal, the effect of small-scale fading can be averaged out, narrow and high-gain beams can be formed, and high spectral efficiency can be achieved [2]–[8].

Despite the great potential of massive MIMO, it is difficult to apply the traditional digital beamforming techniques to a massive MIMO system in practice because one dedicated radio frequency (RF) chain is required for each antenna element, which is of high cost and with high power consumption. To reduce the number of required RF chains in a massive MIMO system, research has been focused on analog beamforming [9]–[12] where a beam is formed by adjusting the independent phase shifters on the BS antennas at the RF part (i.e., only one RF chain is needed for a single data stream), and hybrid analog-digital beamforming [13]–[22].

For both analog and hybrid analog-digital beamforming systems, a fixed-beam network along with beam selection has emerged as a popular technique due to its simplicity [12]–[16]. In this scheme, a fixed number of beams are generated and pointed to different predetermined directions to cover the whole cell, and each selected beam is connected to a dedicated RF chain. By applying this fixed-beam network in hybrid analog-digital beamforming systems, a number of analog beams are first selected and a digital beamformer is then adopted based on the selected analog beams to serve all of the users in the system, which requires that the number of RF chains in the system is no less than the number of users [13]–[16]. A simple fixed-beam based pure analog beamforming system was considered in [12] where each user is served by an allocated beam, implying that the number of RF chains in the system is equal to the number of users. However, as the number of smartphone users is rapidly increasing, the number of RF chains in a practical system will probably be smaller than the number of users. It is therefore of great practical importance to focus on the scenario that the number of RF chains is smaller than the number of users.

In this paper, we first revisit the beam allocation problem investigated in [12] for a fixed-beam based multiuser massive MIMO[1] system where the number of beams N is much larger than the number of users K, i.e., $N \gg K$, by adding a practical constraint that the number of RF chains is smaller than the number of users, implying that the number of simultaneously served users is no larger than the number of RF chains in the system. By assuming universal frequency reuse among the beams, a greedy beam allocation (GBA) algorithm is proposed in this paper to maximize the sum data rate, which achieves nearly the same performance as the optimal brute-force search. However, as the users are randomly located within the cell and have distinct angular separation from the main direction of their serving beams, there are always users, especially those at a serving beam edge, suffering from low power efficiency and high inter-beam interference. These users are defined as "worst-case" users. Since it is equally important to maximize the sum data rate and reduce the rate disparity of the users,

Manuscript received March 6, 2017; revised July 31, 2017; November 28, 2017; accepted December 30, 2017. This paper will be presented in part at the IEEE International Conference on Communications (ICC), Kansas City, May 2018. The associate editor coordinating the review of this paper and approving it for publication was E. Jorswieck.

J. Wang, H. Zhu, N. J. Gomes, and J. Wang are with the School of Engineering and Digital Arts, University of Kent, Canterbury, Kent, CT2 7NT, UK (email: {jw712, h.zhu, n.j.gomes, j.z.wang}@kent.ac.uk).

[1]Massive MIMO is referred to in this paper to indicate generally a system with a large-scale antenna array at the BS, and larger number of antennas than number of users.

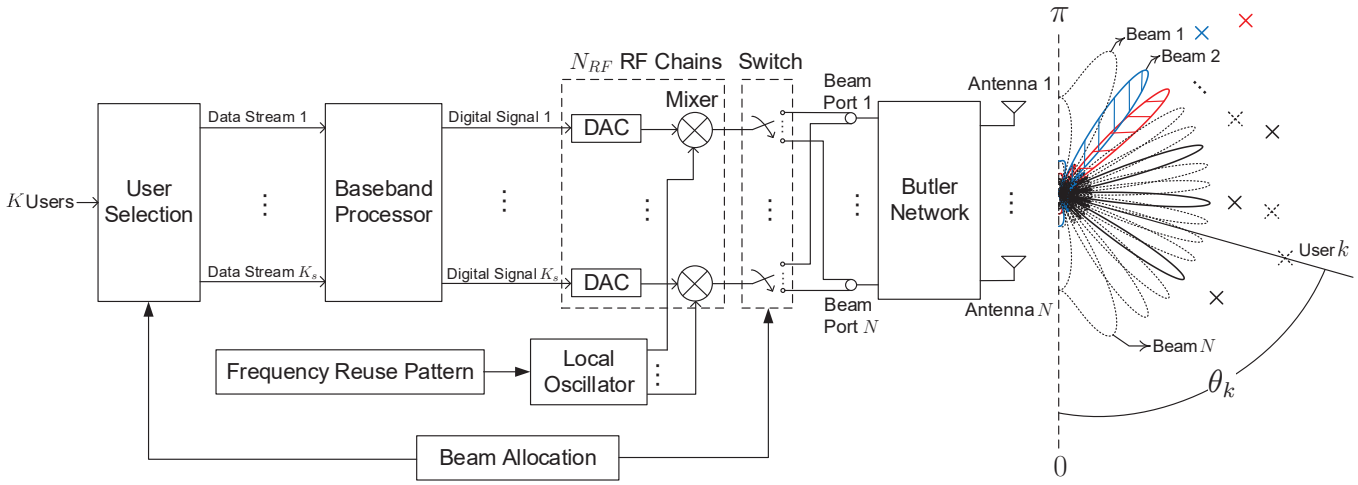


Fig. 1. Illustration of downlink multiuser massive MIMO systems with beam allocation. "x" represents a user. The served users and active beams are plotted in solid lines while the unserved users and inactive beams are plotted in dashed lines. A beam is allocated to the user in the same color.

in this paper, we further aim to improve the worst-case users' data rates.

Note that the low data rate of a worst-case user results from the severe inter-beam interference coming from the adjacent beams of its serving beam. To mitigate the inter-beam interference for the worst-case users, we propose to allocate distinct frequency bands for the adjacent beams. A simple regular fixed frequency reuse scheme is first investigated in our beam allocation based multiuser massive MIMO system, which is shown to be able to improve the data rates of the worst-case users efficiently. However, the sum data rate with fixed frequency reuse degrades significantly compared to universal frequency reuse. An adaptive frequency reuse scheme is therefore further proposed to attain the benefits of both fixed frequency reuse and universal frequency reuse. Simulation results show that our proposed adaptive frequency reuse scheme can improve the worst-case users' data rates and achieve similar sum data rate to universal frequency reuse. Moreover, the proposed adaptive frequency reuse achieves higher average minimum data rate of the served users than both universal frequency reuse and fixed frequency reuse, implying that our proposed adaptive frequency reuse scheme improves the max-min fairness of the served users. In addition, our proposed adaptive frequency reuse scheme only requires users' location information, and more importantly, the frequency band allocations for different users are independent from each other, indicating that frequency bands can be allocated to users in parallel. Therefore, our proposed adaptive frequency reuse scheme is of low complexity and can be easily adopted in future beam allocation based multiuser massive MIMO systems.

The remainder of this paper is organized as follows. Section II introduces the system model. A greedy beam allocation algorithm is proposed in Section III, based on which fixed and adaptive frequency reuse strategies are proposed in Section IV. Discussion is provided in Section V, and concluding remarks are summarized in Section VI.

Throughout this paper, $\mathbb{E}[\cdot]$ denotes the expectation operator. $|X|$ denotes the cardinality of set X. $X \cap Y$ and $X \cup Y$ denote the intersection and union of set X and set Y, respectively. $X \setminus Y$ denotes the relative compliment of set Y in set X. $\emptyset$ denotes the empty set.

II. System Model

Consider the downlink transmission of a beam allocation based multiuser massive MIMO system as shown in Fig. 1. It is assumed that K users are uniformly distributed within a circular cell with unit radius, and each user is equipped with a single antenna. A massive number of N fixed beams are formed by deploying the Butler network [23] with a linear array of N identical isotropic antenna elements at the BS. Denote the set of users as $\mathcal{K}$ and the set of fixed beams as $\mathcal{B}$, with $|\mathcal{K}| = K$ and $|\mathcal{B}| = N$. The BS is located at the center of the cell and all the BS antenna elements are equally spaced at distance $d = 0.5\lambda$, where λ is the propagation wavelength. A fixed number of N_{RF} RF chains are embedded in the system. As employing RF chains is expensive, the number of RF chains N_{RF} is usually not large and assumed to be smaller than the number of users K.

For the beam allocation based system, each user is served by allocating a beam to it. To avoid severe intra-beam interference, one beam is allocated to at most one user. Denote the set of served users as $\mathcal{K}_s$ and the corresponding set of serving beams as $\mathcal{B}_s$ with $|\mathcal{K}_s| = |\mathcal{B}_s| = K_s$. Since each data stream requires an individual RF chain for transmission, the number of simultaneously served users/active beams is limited by the number of RF chains N_{RF}, i.e., $K_s \leq N_{RF}$. Under the practical constraint that N_{RF} is smaller than the total number of users K, not all the users can be served simultaneously. As shown in Fig. 1, with the feedback of the beam allocation result, i.e., which beam is allocated to which user for data transmission, K_s out of K users are selected to be served and their K_s corresponding data streams are sent for baseband waveform processing. Then, each digital basedband output signal passes through its own RF chain where the input digital signal is first converted to an analog signal by using a digital-to-analog converter (DAC) and further upconverted by a local

oscillator (LO) according to the frequency reuse pattern. After upconversion, the RF signal is fed into a particular beam port n via a switch to activate the allocated beam n for its data transmission.

It can be seen from Fig. 1 that the data transmission is determined by the beam allocation and frequency reuse strategies. Specifically, with N beams, K users and N_{RF} RF chains available, beam allocation determines how to efficiently allocate beams to users for data transmission, which provides information for the user selection part and beam switching part, and will be discussed in Section III. Frequency reuse pattern determines the working frequency of each active beam. For example, in Fig. 1, the solid black beams are allocated with the whole frequency band as their adjacent interfering beams are off and the inter-beam interference is small. In this case, using the whole frequency band for transmission can achieve higher data rate. For the shaded blue and red beams, two distinct subbands are allocated since their served users are very close to each other, and the cross-interference is strong and needs to be eliminated. Therefore, it is important to design an efficient frequency reuse scheme to mitigate the inter-beam interference, which will be investigated in Section IV.

By applying the Butler network to form $N = 2^i$ beams where $i \geq 1$ is an integer, the directivity of any beam $n \in \mathcal{B}$, i.e., beam gain achieved relative to an isotropic antenna, with respect to an angle of departure (AoD) θ of the signal is given by [12]

$$D_n(\theta) = \frac{\sin^2(0.5N\pi\cos\theta - \beta_n)}{N\sin^2(0.5\pi\cos\theta - \frac{1}{N}\beta_n)}, \tag{1}$$

where

$$\beta_n = \left(-\frac{N+1}{2} + n\right)\pi. \tag{2}$$

User $k \in \mathcal{K}$ located at (ρ_k, θ_k) is considered as the reference user. By assuming a line-of-sight (LOS) channel at millimeter-wave (mmWave) frequencies, the AoD of the signal received at user k is θ_k as illustrated in Fig. 1, and the corresponding received power of the desired signal can be written as [25]

$$P_k = \sum_{n\in\mathcal{B}} c_{k,n} \cdot p_n \cdot D_n(\theta_k) \cdot \rho_k^{-\alpha}, \tag{3}$$

where p_n denotes the transmit power allocated to beam n. ρ_k is the distance from the cell center to user k and α is the path-loss exponent. In (3), $c_{k,n} \in \{0, 1\}$ denotes the beam allocation indicator. If beam n is allocated to user k, $c_{k,n} = 1$; otherwise, $c_{k,n} = 0$. The number of served users/active beams, K_s, can then be obtained as

$$K_s = \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{B}} c_{k,n}. \tag{4}$$

Assuming that the total transmit power at the BS is fixed at P_t and equally allocated to all the active beams, the transmit power allocated on beam $n \in \mathcal{B}$ is given by

$$p_n = \begin{cases} \frac{P_t}{K_s}, & \text{if} \quad \sum_{k\in\mathcal{K}} c_{k,n} = 1, \\ 0, & \text{if} \quad \sum_{k\in\mathcal{K}} c_{k,n} = 0. \end{cases} \tag{5}$$

As the users are served by allocating individual beams to them, the beam allocation problem will be first studied in the following section.

III. Beam Allocation under Universal Frequency Reuse

Let us first focus on the beam allocation problem with universal frequency reuse. By assuming that the total system bandwidth is normalized to unity, the achievable data rate of user $k \in \mathcal{K}$ can be written as

$$R_k^{uni} = \log_2\left(1 + \frac{P_k}{\sigma_0^2 + I_k^{uni}}\right), \tag{6}$$

where σ_0^2 is the variance of the additive white Gaussian noise (AWGN), and I_k^{uni} is the inter-beam interference power received at user k, given by

$$I_k^{uni} = \sum_{j\in\mathcal{K}, j\neq k} \sum_{n\in\mathcal{B}} c_{j,n} \cdot p_n \cdot D_n(\theta_k) \cdot \rho_k^{-\alpha}. \tag{7}$$

In this section, we aim at maximizing the sum data rate of the system by properly allocating beams to users when the number of RF chains N_{RF} is fixed and smaller than the total number of users K. The optimization problem will be formulated in the following subsection.

A. Problem Formulation

With the number of RF chains fixed at N_{RF}, the beam allocation problem for maximizing the sum data rate of the multiuser massive MIMO system can be formulated as

$$\max_{\{c_{k,n}\}_{k\in\mathcal{K}, n\in\mathcal{B}}} \quad \sum_{k\in\mathcal{K}} R_k^{uni} \tag{8a}$$

$$\text{subject to} \quad \sum_{n\in\mathcal{B}} c_{k,n} \leq 1, \forall k \in \mathcal{K}; \sum_{k\in\mathcal{K}} c_{k,n} \leq 1, \forall n \in \mathcal{B};$$

$$c_{k,n} \in \{0, 1\}, \forall k \in \mathcal{K}, \forall n \in \mathcal{B}, \tag{8b}$$

$$\sum_{n\in\mathcal{B}} \sum_{k\in\mathcal{K}} c_{k,n} \leq N_{RF}, \tag{8c}$$

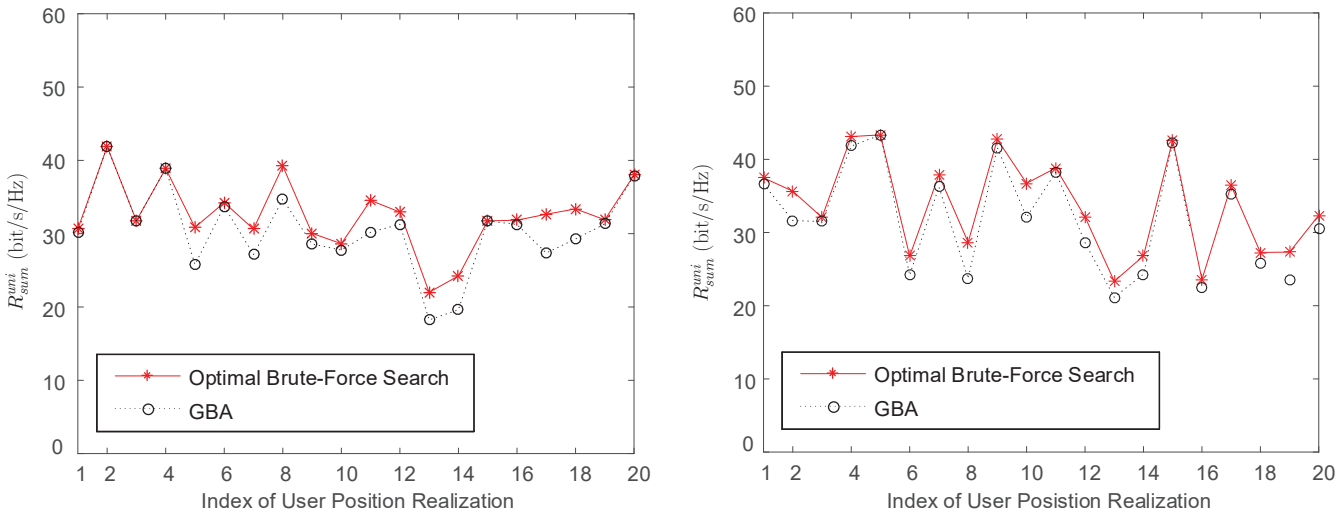
where (8b) is based on the assumptions that each user is served by allocating a beam to it, and each beam can be used by at most one user to avoid intra-beam interference. (8c) follows the constraint that the total number of served users cannot exceed the total number of RF chains N_{RF} since each data stream requires a single RF chain for transmission.

B. Greedy Beam Allocation (GBA)

Note that a special case of the above beam allocation problem given in (8a)–(8c) has been studied in [12] where the number of RF chains N_{RF} is the same as the number of users K. In this special case, the constraint given in (8c) is always satisfied and thus can be removed. Then the beam allocation problem given in (8a)–(8b) is solved by decomposing the beam allocation problem into two subproblems, including a beam-user association subproblem and a beam allocation subproblem. For the beam-user association problem, each user k is associated with beam $n_k^{(1)}$ with the largest directivity, i.e.,

$$n_k^{(1)} = \arg\max_{n\in\mathcal{B}} D_n(\theta_k). \tag{9}$$

The beam allocation subproblem is then reformulated as a monotone submodular maximization problem subject to a



(a) $N_{RF} = 6$ (b) $N_{RF} = 8$

Fig. 2. Achievable sum data rate R_{sum}^{uni} under 20 realizations of users' positions with both the optimal brute-force search and the proposed GBA algorithm. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $N = 16$, $K = 10$.

single partition matroid constraint, which can be efficiently solved by a greedy algorithm.

In this paper, by taking into account that the number of RF chains N_{RF} in practical systems is smaller than the number of users K due to the high cost of deploying RF chains, (8c) cannot be removed. Nevertheless, it can be easily shown that the constraint given in (8c) forms a uniform matroid constraint. Specifically, let us first define a ground set U as

$$U = \{u_{k,n_k^{(1)}} : k \in \mathcal{K}\}, \tag{10}$$

and the beam allocation set S as a subset of U such that $u_{k,n_k^{(1)}} \in S$ if beam $n_k^{(1)}$ is allocated to user k, i.e., $c_{k,n_k^{(1)}} = 1$; otherwise, $u_{k,n_k^{(1)}} \notin S$. The constraint given in (8c) can be then rewritten as

$$S \in \mathcal{I}_{RF}, \tag{11}$$

where

$$\mathcal{I}_{RF} = \{X \subseteq U : |X| \leq N_{RF}\}, \tag{12}$$

which is a uniform matroid. The definition of uniform matroids can be found in Appendix A.

As constraint (8c) is a matroid constraint, the sum data rate maximization problem given by (8a)–(8c) can be transformed into two solvable subproblems by following the submodular maximization method in [12]. Specifically, each user is first associated with its best beam $n_k^{(1)}$ with the largest directivity. For all the associated beam-user pairs, a greedy algorithm can be then employed to efficiently allocate beams to users which selects the best beam-user pair with the highest received signal power at each step, i.e, allocating the selected beam to its associated user, until the constraints are no longer satisfied. This algorithm is referred to as *greedy beam allocation (GBA)* in the following, and the description is presented as Algorithm 1.

Algorithm 1 Greedy Beam Allocation (GBA)

1: **Initialization:** $\mathcal{K}_s = \emptyset$. $\mathcal{K}^{'} = \mathcal{K}$. $c_{k,n} = 0, \forall k \in \mathcal{K}, \forall n \in \mathcal{N}$.
2: **for** $k \in \mathcal{K}$ **do**
3: $\quad n_k^{(1)} = \arg\max_{n \in \mathcal{B}} D_n(\theta_k)$;
4: **end for**
5: **while** $|\mathcal{K}_s| \leq N_{RF}$ & $\mathcal{K}^{'} \neq \emptyset$ **do**
6: $\quad k^* = \arg\max_{k \in \mathcal{K}^{'}} D_{n_k^{(1)}}(\theta_k) \cdot \rho_k^{-\alpha}$,
7: $\quad c_{k^*,n_{k^*}^{(1)}} = 1$, $\mathcal{K}^{'} = \mathcal{K}^{'} \setminus \{j\}_{n_j^{(1)} = n_{k^*}^{(1)}, j \in \mathcal{K}^{'}}$;
8: **end while**

It is clear from Algorithm 1 that in the first beam-user association part, there are K iterations. For each iteration, $\log_2 N$ comparisons are needed to find the best beam of the user by adopting binary search. In the second greedy beam allocation part, as there are totally K beam-user pairs, the total number of comparisons required to find maximum N_{RF} best pairs is $N_{RF}(2K - N_{RF} - 1)/2$. As a consequence, the maximum number of comparisons required by our GBA algorithm is $K \log_2 N + N_{RF}(2K - N_{RF} - 1)/2$.

C. Simulation Results

Fig. 2 presents the simulation results of the achievable sum data rate, $R_{sum}^{uni} \triangleq \sum_{k \in \mathcal{K}} R_k^{uni}$, under 20 random realizations of users' positions when the number of beams $N = 16$, the number of users $K = 10$, and the number of RF chains $N_{RF} = 6$ and 8, respectively.[2] It can be seen from Fig. 2 that our proposed GBA algorithm can achieve sum data rate very close to that of the optimal brute-force search under most realizations.

[2]As the computational complexity of brute-force search $O(N^K)$ sharply increases with the number of beams N and the number of users K, simulation results with $N = 16$, $K = 10$, and the number of RF chains $N_{RF} = 6$ and 8 are presented.

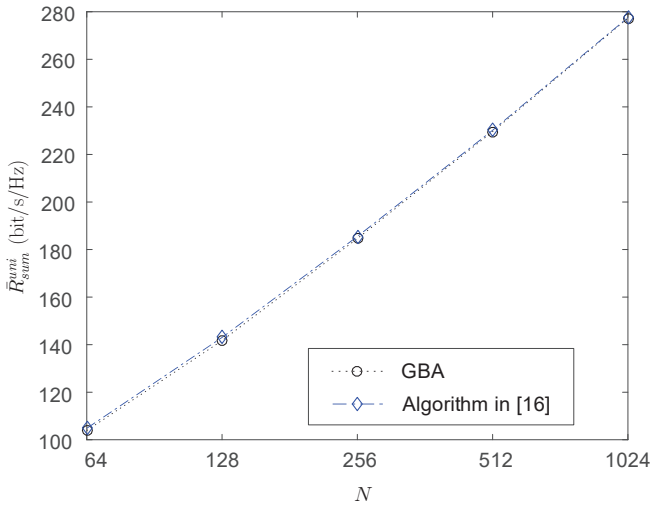


Fig. 3. Average achievable data rate $\bar{R}_{sum}^{uni}$ with the proposed GBA algorithm and the algorithm in [16]. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $K = 30$, $N_{RF} = 30$.

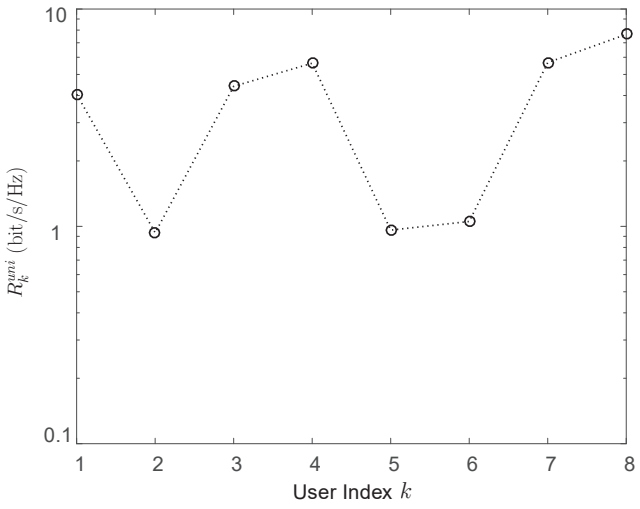


Fig. 5. Achievable data rate R_k^{uni} of each served user $k \in \mathcal{K}_s$ with the proposed GBA algorithm under the topology given in Fig. 4. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $N = 16$, $K = 10$, $N_{RF} = 8$.

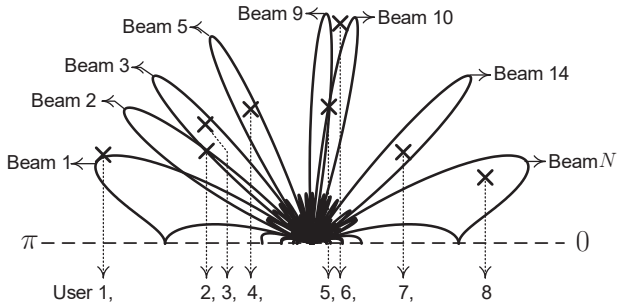


Fig. 4. Illustration of beam allocation result for a random realization of users' positions with the proposed GBA algorithm. "x" represents a served user. $\alpha = 2.2$, $N = 16$, $K = 10$, $N_{RF} = 8$.

Fig. 3 further compares the average sum data rate $\bar{R}_{sum}^{uni} \triangleq \mathbb{E}_{\{\mathbf{r}_k:\ k\in\mathcal{K}\}}[\sum_{k\in\mathcal{K}} R_k^{uni}]$ achieved by our proposed GBA algorithm with that achieved by the algorithm in [16]. To make a fair comparison, the number of RF chains N_{RF} is set to be the same as the number of users K as this is required by the algorithm in [16]. It is shown in Fig. 3 that our algorithm achieves almost the same average sum data rate as the algorithm in [16] which has the computational complexity of $O(K^2N)$. For our proposed GBA algorithm, with $N_{RF} = K$, it reduces to the low-complexity beam allocation algorithm in [12] which has the computational complexity of $O(K \log_2 N)$. It can be clearly seen that while providing the same performance, our GBA algorithm has a much lower computational complexity than the algorithm in [16], especially in a massive MIMO system where the number of beams N and the number of users K are both large.

To take a close look at the individual data rates of the served users, Fig. 4 illustrates the beam allocation result by applying the GBA algorithm to a random realization of users' positions where only the active beams are drawn and each active beam serves the user falling into its own angular coverage.[3] It can be seen from Fig. 4 that there are a few "beam-edge" users. A beam-edge user is close to the angular edge between its serving beam and an adjacent beam which is active for serving another user. For example, user 2 is served by beam 2 and located at the angular edge between beam 2 and beam 3. As beam 3 is allocated to user 3 for data transmission, user 2 suffers from strong inter-beam interference from beam 3 and thus achieves low data rate which can be observed from Fig. 5. Similarly, user 5 and user 6 achieve very low data rate due to the strong inter-beam interference from each other. In particular, the achievable data rates of user 2 and user 5 are even lower than 1 bit/s/Hz, implying that the interference power is even higher than the desired signal power. In this paper, these beam-edge users are referred to as *worst-case users*, which will be defined in the following subsection.

D. Definition of Worst-Case Users

It is clear from the above discussion that a user $k \in \mathcal{K}_s$ is a worst-case user if two conditions are satisfied: (1) its strongest potential interfering beam is active for serving another user; and (2) it is close to the angular edge between its serving beam and strongest potential interfering beam. For the first condition, given any served user k located at (ρ_k, θ_k), let

$$D_{n_k^{(1)}}(\theta_k) \geq D_{n_k^{(2)}}(\theta_k) \geq \cdots \geq D_{n_k^{(N)}}(\theta_k) \qquad (13)$$

denote the order statistics obtained by arranging the directivities $D_1(\theta_k), D_2(\theta_k), \cdots, D_N(\theta_k)$ of N beams with respect to user k, where $n_k^{(l)}$ denotes the lth best beam of user k with the lth largest directivity. Particularly, beam $n_k^{(1)}$ is the beam allocated to user k for data transmission, and beam $n_k^{(2)}$ is the strongest potential interfering beam. Therefore, if a user is a worst-case user, beam $n_k^{(2)}$ must be active.

[3]To clearly see the beam pattern and users' locations, a beam allocation result for a random realization of 10 users' positions with 16 beams is presented.

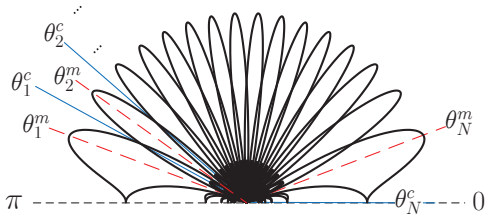


Fig. 6. Illustration of the main direction of beam n, θ_n^m, and the direction of the angular edge between beam n and beam $n+1$, θ_n^c.

For the second condition, to quantify the closeness of a served user k to the angular edge between beam $n_k^{(1)}$ and beam $n_k^{(2)}$, let us first define a general mapping as

$$\psi = \cos\theta, \tag{14}$$

where θ is the angular coordinate. For the Butler network with N beams formed, as illustrated in Fig. 6, θ_n^m denotes the main direction of beam n and θ_n^c denotes the direction of the angular edge between beam n and beam $n+1$. According to Appendix B, for any beam $n \in \mathcal{B}$, we have

$$\psi_n^m = \cos\theta_n^m = \frac{2n-1}{N} - 1, \tag{15}$$

from (B.4) and

$$\psi_n^c = \cos\theta_n^c = \frac{2n}{N} - 1, \tag{16}$$

from (B.7), respectively. It can be then easily obtained from (15)–(16) that

$$\Delta\psi = \psi_n^c - \psi_n^m = \psi_n^m - \psi_{n-1}^c = \frac{1}{N}, \tag{17}$$

for any $n \in \mathcal{B}$, which is independent of the beam index n. This indicates that (1) every beam has the same angular coverage in the domain of ψ, and (2) the angular coverage of each beam n is symmetric to its main direction ψ_n^m as shown in Fig. 7. Therefore, the closeness from user k at (ρ_k, θ_k) to the angular edge between beam $n_k^{(1)}$ and beam $n_k^{(2)}$ can be defined as

$$\Delta\psi_k = |\psi_k - \psi_{\tilde{n}_k}^c|, \tag{18}$$

where $\psi_k = \cos\theta_k$ and $\psi_{\tilde{n}_k}^c$ denotes the direction of the angular edge between beam $n_k^{(1)}$ and beam $n_k^{(2)}$ with index $\tilde{n}_k$ given by

$$\tilde{n}_k = \min\{n_k^{(1)}, n_k^{(2)}\}. \tag{19}$$

By combining (17)–(19), we have

$$\Delta\psi_k \le \Delta\psi = \frac{1}{N}. \tag{20}$$

Intuitively, a smaller $\Delta\psi_k$ indicates that user k is closer to its strongest potential interfering beam. Therefore, if a user is a worst-case user, the angular separation $\Delta\psi_k$ must be smaller than a given threshold $\Delta\psi_{th}$. To conclude, the definition of worst-case users is given as follows.

Definition 1. *A user $k \in \mathcal{K}$ is a worst-case user, denoted by $k \in \mathcal{K}_{worst}$, with the following properties:*
(1) User k is served, i.e., $k \in \mathcal{K}_s$;

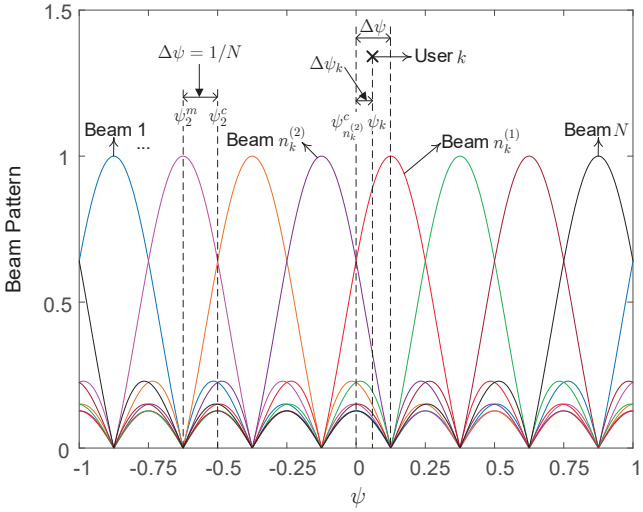


Fig. 7. Illustration of beam pattern in the domain of $\psi = \cos\theta$. $N = 8$.

(2) User k's strongest potential interfering beam is allocated to another user for data transmission, i.e., beam $n_k^{(2)}$ is active;
(3) The angular separation from user k to the edge between beam $n_k^{(1)}$ and beam $n_k^{(2)}$, $\Delta\psi_k$, is no larger than the threshold $\Delta\psi_{th}$, i.e., $\Delta\psi_k \le \Delta\psi_{th}$.

As the worst-case users could suffer from very strong inter-beam interference and thus very low data rate, we are particularly interested in improving the data rates of the worst-case users in this paper. To this end, frequency reuse will be introduced in the following section to mitigate the strong inter-beam interference.

IV. Frequency Reuse of Greedy Beam Allocation

A. Fixed Frequency Reuse

As described in Section III, although the proposed GBA algorithm can achieve near-optimal sum data rate, there could be some worst-case users who suffer from very low data rate due to strong inter-beam interference. To improve the data rates of these users, we propose to allocate different frequency bands to the adjacent beams for data transmission so that the interference from adjacent beams, which contributes to most of the inter-beam interference, can be eliminated. A simple method is the fixed frequency reuse. Specifically, for a frequency reuse factor $f < 1$, the whole frequency band is equally divided into $1/f$ subbands. These subbands are then assigned to the beams from beam 1 to beam N, i.e., from the left-hand side to the right-hand side, regularly as Fig. 8 illustrates. Note that beam 1 and beam N can be seen as two adjacent beams since the largest sidelobe of beam 1 (beam N) is overlapped with the main lobe of beam N (beam 1). As the number of beams N is a power of 2 in a Butler network, to ensure that any two adjacent beams have different frequency bands allocated, the inverse of the frequency reuse factor, $1/f$, must be even. A counter example is that if $f = 1/3$ with $N = 16$, both beam 1 and beam 16 use subband 1 for data transmission which causes strong inter-beam interference.

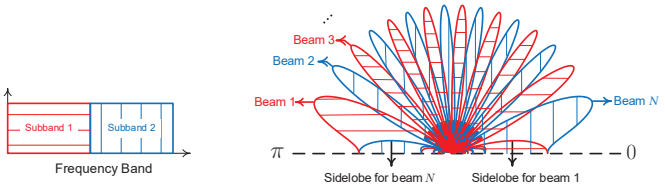


Fig. 8. Illustration of fixed frequency reuse. $N = 16$, $f = 1/2$.

With fixed frequency reuse factor f, let $\mathcal{K}_i^{fix}$ denote the set of users using subband i for data transmission with $\bigcup_{i=1,2,\cdots,1/f} \mathcal{K}_i^{fix} = \mathcal{K}_s$. The achievable data rate of user $k \in \mathcal{K}_i^{fix}$ can be then obtained as

$$R_k^{fix} = f \log_2 \left(1 + \frac{P_k}{f\sigma_0^2 + I_k^{fix}}\right), \tag{21}$$

where the inter-beam interference power received at user k is given by

$$I_k^{fix} = \sum_{j \in \mathcal{K}_i^{fix}, j \neq k} p_{n_j^{(1)}} \cdot D_{n_j^{(1)}}(\theta_k) \cdot \rho_k^{-\alpha}. \tag{22}$$

Fig. 9 presents the achievable data rate R_k^{fix} of each served user $k \in \mathcal{K}_s$ with the fixed frequency reuse factor $f = 1/2$. For the sake of comparison, the achievable data rate R_k^{uni} with universal frequency reuse, i.e., $f = 1$, shown in Fig. 5 is also presented. It can be seen from Fig. 9 that with fixed frequency reuse factor $f = 1/2$, the data rates of low-data-rate users 2, 5 and 6 under universal frequency reuse are improved, thanks to the elimination of the dominant interference coming from the adjacent beams. However, the data rates of the other users are reduced significantly compared to those with universal frequency reuse due to only half the bandwidth being allocated, which results in a sacrifice in the sum data rate.

As the data rate R_k^{fix} closely depends on the locations of users $\{\mathbf{r}_k : k \in \mathcal{K}\}$, to assess the effect of the frequency reuse factor f, we further focus on the average achievable data rate of a served user k given $\Delta\psi_k$, which is defined as $\bar{R}_k^{fix} \triangleq \mathbb{E}_{\{\mathbf{r}_k:\ k\in\mathcal{K}\}}\left[R_k^{fix} \mid k \in \mathcal{K}_s, \Delta\psi_k\right]$, and the average sum data rate, defined as $\bar{R}_{sum}^{fix} \triangleq \mathbb{E}_{\{\mathbf{r}_k:\ k\in\mathcal{K}\}}\left[\sum_{k\in\mathcal{K}_s} R_k^{fix}\right]$.

Fig. 10(a) presents the average achievable data rate of user $k \in \mathcal{K}_s$ whose strongest potential interfering beam $n_k^{(2)}$ is active versus the angular separation $\Delta\psi_k$ from user k to the edge between beam $n_k^{(2)}$ and its serving beam $n_k^{(1)}$ under varying frequency reuse factor f. It is clearly seen from Fig. 10(a) that the average achievable data rate $\bar{R}_k^{fix}$ increases as the frequency reuse factor increases from $f = 1/8$ to $f = 1/2$. The reason is that with a low f, although the inter-beam interference is effectively eliminated, the frequency band allocated for each beam is narrow, which results in low data rate. As frequency reuse factor f increases, the width of the allocated frequency band for each beam increases and thus the data rate increases. For our considered linear antenna array, the strongest inter-beam interference for user k comes from its adjacent beams. As a result, with a small $\Delta\psi_k$, i.e., the user is close to its strongest interfering beam, the optimal

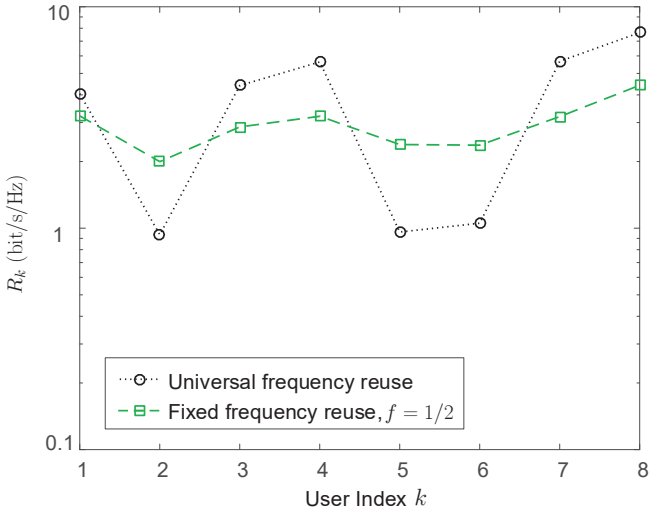


Fig. 9. Achievable data rate R_k of each served user $k \in \mathcal{K}_s$ with fixed frequency reuse factor $f = 1/2$ and universal frequency reuse under the topology given in Fig. 4. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $N = 16$, $K = 10$, $N_{RF} = 8$.

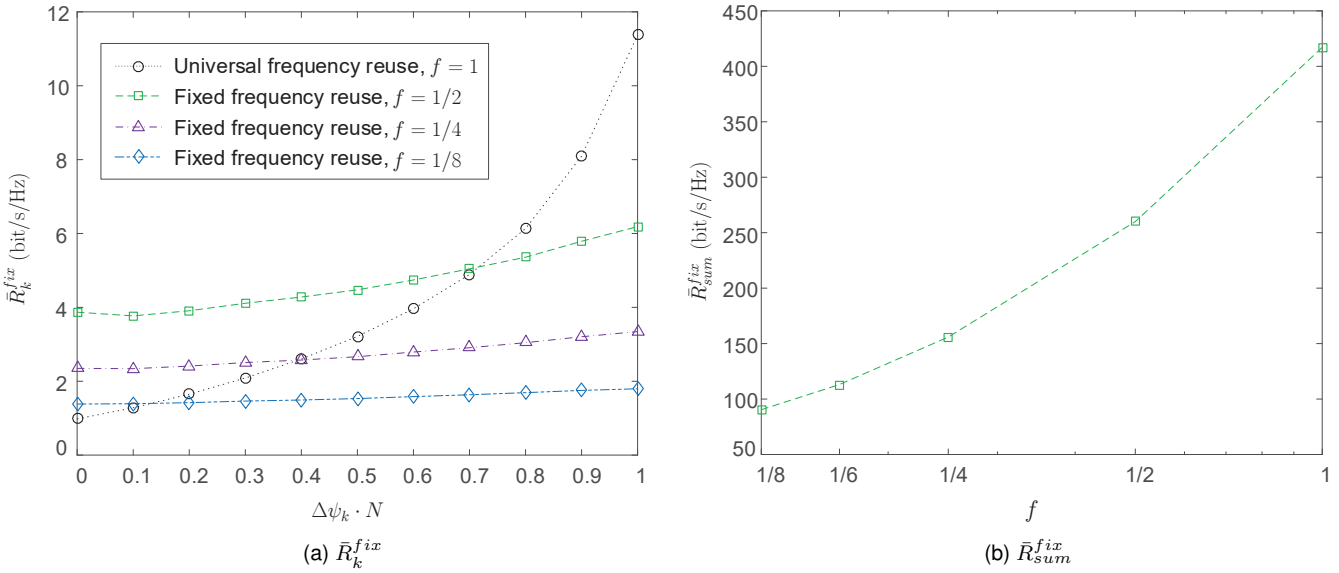
frequency reuse factor to maximize its average data rate is $f = 1/2$. By contrast, when the user is close to the main direction of its serving beam, i.e., $\Delta\psi_k$ is large and approaches $1/N$, the inter-beam interference from its strongest interfering beam becomes very small and thus using universal frequency reuse can maximize its average data rate.

We can then conclude that with fixed frequency reuse factor $f = 1/2$, the achievable data rates of the users who are severely interfered under universal frequency reuse can be improved while the data rates of the users with low inter-beam interference under universal frequency reuse are significantly reduced. As the number of served users is smaller than the number of beams in a massive MIMO system, only some of the users suffer from very strong inter-beam interference. Therefore, by using a fixed frequency reuse factor $f = 1/2$ in the system, the average sum data rate is much lower than that with universal frequency reuse as shown in Fig. 10(b). Since the optimal frequency reuse factor for a served user varies with the layout of the active beams, and the average sum data rate performance is significantly decreased by using fixed frequency reuse, an adaptive frequency reuse scheme will be proposed in the following subsection.

B. Adaptive Frequency Reuse

As can be seen from Figs. 9–10, a user k who is close to its active strongest interfering beam $n_k^{(2)}$ achieves the highest average data rate with frequency reuse factor $f = 1/2$; otherwise it achieves the highest average data rate with frequency reuse factor $f = 1$. An adaptive frequency reuse strategy is then proposed as follows.

For any worst-case user $k \in \mathcal{K}_{worst}$, to eliminate the strong inter-beam interference from beam $n_k^{(2)}$, we propose to allocate distinct half frequency bands to its serving beam $n_k^{(1)}$ and interfering beam $n_k^{(2)}$ for data transmission. Specifically, let



(a) $\bar{R}_k^{fix}$ (b) $\bar{R}_{sum}^{fix}$

Fig. 10. (a) Average achievable data rate $\bar{R}_k^{fix}$ of each user $k \in \mathcal{K}_s$ whose strongest potential interfering beam $n_k^{(2)}$ is active versus the angular separation $\Delta\psi_k$ with fixed frequency reuse. (b) Average sum data rate $\bar{R}_{sum}^{fix}$ versus fixed frequency reuse factor f. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $N = 512$, $K = 80$, $N_{RF} = 60$.

$\mathcal{B}_{half}^{adp}$ denote the corresponding set of the worst-case users' serving beams and strongest interfering beams, i.e., $\mathcal{B}_{half}^{adp} = \{n_k^{(1)}, n_k^{(2)} : k \in \mathcal{K}_{worst}\}$. For each beam $n \in \mathcal{B}_{half}^{adp}$, if n is odd, the first half frequency band (Subband 1) is allocated, i.e., $n \in \mathcal{B}_1^{adp}$; if n is even, the second half frequency band (Subband 2) is allocated, i.e., $n \in \mathcal{B}_2^{adp}$. For the rest of active beams, full frequency band is allocated. The corresponding beam set is denoted by $\mathcal{B}_{full}^{adp}$ with $\mathcal{B}_{full}^{adp} = \mathcal{B}_s \setminus \mathcal{B}_{half}^{adp}$. The detailed description is presented as Algorithm 2.

As whether a user is a worst-case user closely depends on the threshold $\Delta\psi_{th}$, the adaptive frequency band allocation result is determined by the threshold $\Delta\psi_{th}$. Specifically, with a small $\Delta\psi_{th}$, some users close to the strongest interfering beams may not be included into the worst-case user set and thus full frequency band is allocated which results in strong inter-beam interference and low data rate. While $\Delta\psi_{th}$ is large, some users relatively far away from their strongest interfering beams may be labeled as worst-case users with half frequency band allocated which results in low data rate as well. Therefore, the threshold $\Delta\psi_{th}$ needs to be carefully chosen for improving the individual data rate.

As we aim at improving the achievable data rates of the low-data-rate users with universal frequency reuse by adopting the proposed adaptive frequency reuse algorithm, all the users who can achieve higher data rate by using half frequency band than by using full frequency band should be included in the worst-case user set $\mathcal{K}_{worst}$. As a result, the threshold $\Delta\psi_{th}$ should be maximized as far as possible. That is,

$$\begin{aligned} &\max && \Delta\psi_{th} && \text{(23a)} \\ &\text{subject to} && R_k^{adp} \geq R_k^{uni},\ \forall k \in \mathcal{K}_{worst}, && \text{(23b)} \end{aligned}$$

where R_k^{adp} denotes the achievable data rate of user k with the proposed adaptive frequency reuse scheme which is given by

$$R_k^{adp} = \begin{cases} \log_2\left(1 + \frac{P_k}{\sigma_0^2 + I_k^{adp}}\right), & \text{if} \quad n_k^{(1)} \in \mathcal{B}_{full}^{adp}, \\ \frac{1}{2}\log_2\left(1 + \frac{P_k}{\frac{1}{2}\sigma_0^2 + I_k^{adp}}\right), & \text{if} \quad n_k^{(1)} \in \mathcal{B}_i^{adp}, i = 1, 2. \end{cases} \tag{24}$$

For any user k, if its allocated beam $n_k^{(1)}$ transmits over the whole frequency band, it receives the whole inter-beam

Algorithm 2 Adaptive Frequency Reuse

Input: $\Delta\psi_{th}$, $\mathcal{K}_s$, $\mathcal{B}_s$.
Initialization: $\mathcal{B}_1^{adp} = \emptyset$, $\mathcal{B}_2^{adp} = \emptyset$, $\mathcal{B}_{full}^{adp} = \emptyset$.
for $k \in \mathcal{K}_s$ **do**
  $n_k^{(1)} = \arg\max_{n \in \mathcal{B}} D_n(\theta_k)$, $n_k^{(2)} = \arg\max_{n \in \mathcal{B}, n \neq n_k^{(1)}} D_n(\theta_k)$;
  if $n_k^{(2)} \in \mathcal{B}_s$ **then**
    $\tilde{n}_k = \min\{n_k^{(1)}, n_k^{(2)}\}$, $\Delta\psi_k = |\psi_k - \psi_{\tilde{n}_k}^c|$;
    if $\Delta\psi_k \leq \Delta\psi_{th}$ **then**
      if $n_k^{(1)}$ is odd **then**
        $\mathcal{B}_1^{adp} = \mathcal{B}_1^{adp} \cup \left\{n_k^{(1)}\right\}$, $\mathcal{B}_2^{adp} = \mathcal{B}_2^{adp} \cup \left\{n_k^{(2)}\right\}$;
      else
        $\mathcal{B}_2^{adp} = \mathcal{B}_2^{adp} \cup \left\{n_k^{(1)}\right\}$, $\mathcal{B}_1^{adp} = \mathcal{B}_1^{adp} \cup \left\{n_k^{(2)}\right\}$;
      end if
    end if
  end if
end for
$\mathcal{B}_{full}^{adp} = \mathcal{B}_s \setminus \mathcal{B}_1^{adp} \setminus \mathcal{B}_2^{adp}$.
Output: $\mathcal{B}_1^{adp}$, $\mathcal{B}_2^{adp}$, $\mathcal{B}_{full}^{adp}$.

$$I_k^{adp} = \begin{cases} \sum_{n\in\mathcal{B}_s, n\neq n_k^{(1)}} p_n \cdot D_n(\theta_k) \cdot \rho_k^{-\alpha}, & \text{if} \quad n_k^{(1)} \in \mathcal{B}_{full}^{adp}, \\ \sum_{n\in\mathcal{B}_i^{adp}, n\neq n_k^{(1)}} p_n \cdot D_n(\theta_k) \cdot \rho_k^{-\alpha} + \frac{1}{2}\sum_{n\in\mathcal{B}_{full}^{adp}} p_n \cdot D_n(\theta_k) \cdot \rho_k^{-\alpha}, & \text{if} \quad n_k^{(1)} \in \mathcal{B}_i^{adp}, i = 1, 2. \end{cases} \tag{25}$$

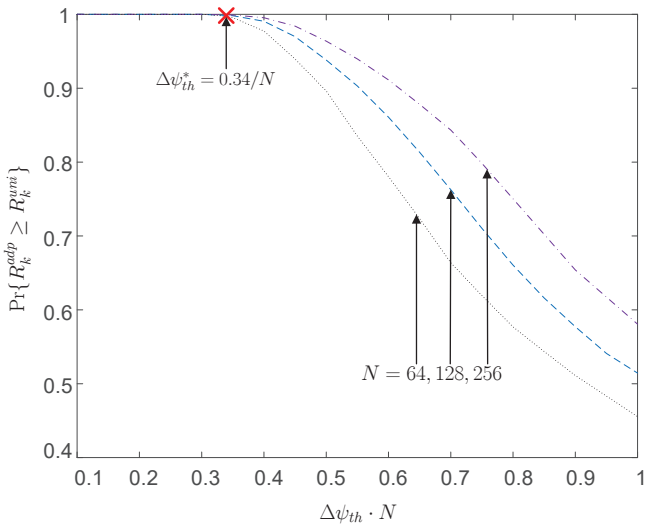


Fig. 11. Probability $\Pr\{R_k^{adp} \geq R_k^{uni}\}$ that the achievable data rate of any worst-case user $k \in \mathcal{K}_{worst}$ with the proposed adaptive frequency reuse is no less than that with universal frequency reuse versus the threshold $\Delta\psi_{th}$. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $K = 30$, $N_{RF} = 20$.

interference from all the active beams. If $n_k^{(1)}$ transmits over subband i, it receives the whole interference from the beams working on the same subband i and half of the interference from the beams working on the full frequency band. Therefore, the inter-beam interference power received at user k, I_k^{adp}, can be obtained as (25) shown at the top of this page.

Appendix C shows that the solution of the optimization problem defined in (23a)–(23b) is approximately obtained as

$$\Delta\psi_{th}^* \approx \left(\sqrt{\frac{2}{C} - C^2 - 1} - C\right) \cdot \frac{1}{N}, \tag{26}$$

where

$$C = \sqrt{\frac{1}{3}\left(\left(17 + 3\sqrt{33}\right)^{\frac{1}{3}} - 2\left(17 + 3\sqrt{33}\right)^{-\frac{1}{3}} - 1\right)} \approx 0.737. \tag{27}$$

By substituting (27) into (26), we have

$$\Delta\psi_{th}^* \approx 0.34\frac{1}{N}. \tag{28}$$

Fig. 11 presents the probability $\Pr\{R_k^{adp} \geq R_k^{uni}\}$ that the achievable data rate of any worst-case user $k \in \mathcal{K}_{worst}$ with adaptive frequency reuse is no less than that with universal frequency reuse. It can be seen from Fig. 11 that with a small threshold $\Delta\psi_{th}$, the probability $\Pr\{R_k^{adp} \geq R_k^{uni}\}$ is 1, implying that all the worst-case users achieve higher data rate by using adaptive frequency reuse. As $\Delta\psi_{th}$ increases, $\Pr\{R_k^{adp} \geq R_k^{uni}\}$ decreases rapidly which indicates that some users who can achieve better data rate performance by using full frequency band instead of half frequency band are included in the worst-case user set. Therefore, to improve the individual data rate by adopting adaptive frequency reuse, the threshold $\Delta\psi_{th}$ should be properly chosen. The optimal threshold $\Delta\psi_{th}^*$ has been derived in (28) which is $0.34/N$. Fig. 11 corroborates that (28) serves as a good approximation of the optimal threshold under varying number of beams N. Thanks to the closed-form optimal threshold in (28), only three simple conditions need to be checked to determine whether a user is a worst-case user and allocate a frequency band to it accordingly, which indicates that our proposed adaptive frequency reuse scheme can be easily applied to future beam allocation based massive MIMO systems due to its simplicity.

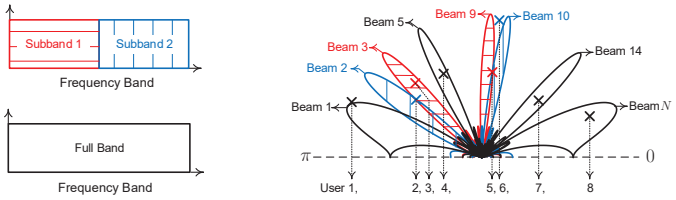


Fig. 12. Illustration of frequency band allocation result by applying the proposed adaptive frequency reuse scheme with optimal threshold $\Delta\psi_{th}^* = 0.34/N$ to the same topology as Fig. 4. $N = 16$, $K = 10$, $N_{RF} = 8$.

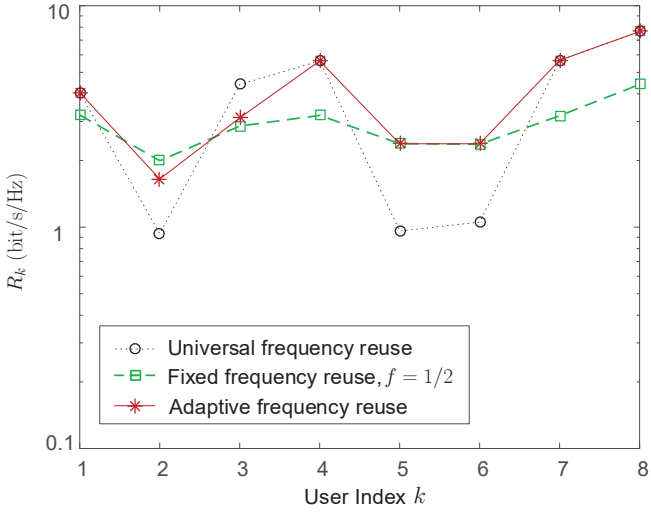
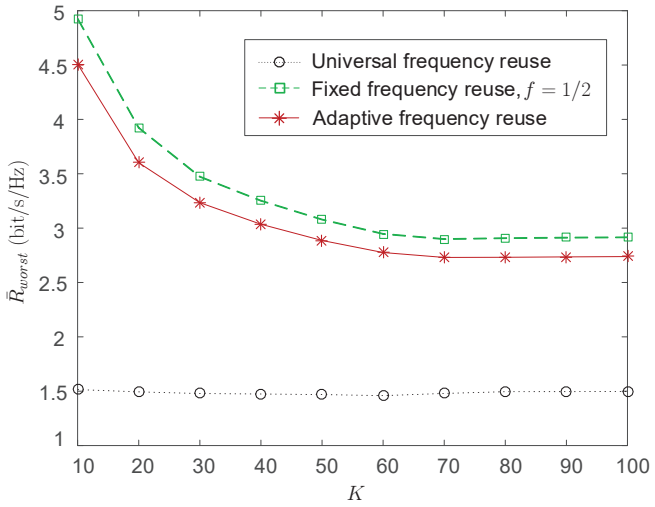
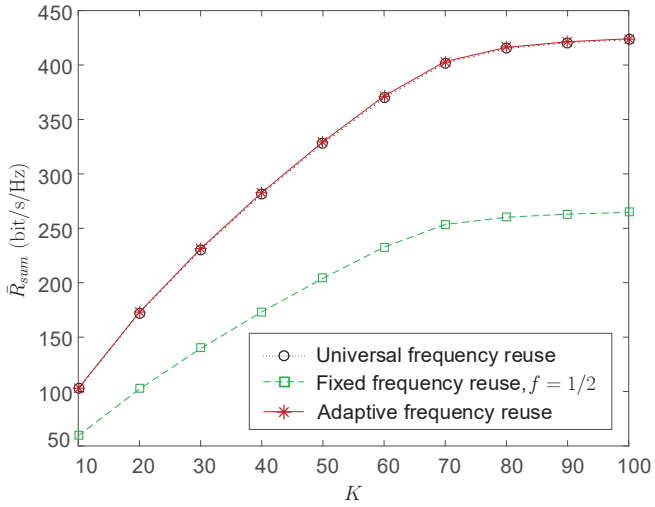


Fig. 13. Achievable data rate R_k of each served user $k \in \mathcal{K}_s$ with the proposed adaptive frequency reuse scheme with $\Delta\psi_{th} = \Delta\psi_{th}^*$, universal frequency reuse and fixed frequency reuse under the topology given in Fig. 12. $\alpha = 2.2$, $P_t/\sigma_0^2 = 20$dB, $N = 16$, $K = 10$, $N_{RF} = 8$.

Fig. 12 presents the frequency band allocation result by applying the proposed adaptive frequency reuse scheme to the topology given in Fig. 4 with the optimal threshold $\Delta\psi_{th}^* = 0.34/N$, where each active beam is allocated with the frequency band in the same color. It can be seen from Fig. 12



(a) $\bar{R}_{worst}$



(b) $\bar{R}_{sum}$

Fig. 14. (a) Average data rate per worst-case user $\bar{R}_{worst}$ and (b) average sum data rate $\bar{R}_{sum}$ with the proposed adaptive frequency reuse, universal frequency reuse and fixed frequency reuse with $f = 1/2$. $\alpha = 2.2$, $\bar{P}_t/\sigma_0^2 = 20$dB, $N = 512$, $N_{RF} = 60$.

that users 2, 5 and 6 are labeled as worst-case users according to Definition 1, i.e., each of them is close to the angular edge between its serving beam and its strongest potential interfering beam and that strongest potential interfering beam is active for another user's data transmission. As a result, half frequency bands are allocated to the serving beams and the strongest interfering beams of users 2, 5 and 6 according to the proposed adaptive frequency reuse scheme.

With the optimal threshold $\Delta\psi_{th}^*$ given in (28), Fig. 13 presents the achievable data rate R_k^{adp} of each served user $k \in \mathcal{K}_s$ with the proposed adaptive frequency reuse. For the sake of comparison, the achievable data rate R_k^{uni} with universal frequency reuse and achievable data rate R_k^{fix} with fixed frequency reuse factor $f = 1/2$ shown in Fig. 9 are also presented. It can be seen from Fig. 13 that with adaptive frequency reuse, the data rates of worst-case users, 2, 5 and 6 are highly improved compared to those with universal frequency reuse and close to those with fixed frequency reuse factor $f = 1/2$, thanks to the mitigation of the dominant interference coming from the strongest interfering beams.

For the remaining users, the proposed adaptive frequency reuse achieves almost the same data rates as universal frequency reuse except user 3. For user 3, since its allocated beam 3 is the strongest interferer of worst-case user 2, half frequency band is allocated to beam 3. Therefore, compared to universal frequency reuse, user 3's data rate is reduced which results in slight sum data rate degradation. It can be concluded from this figure that the proposed adaptive frequency reuse scheme can achieve the benefits of both fixed frequency reuse with $f = 1/2$ and universal frequency reuse. Specifically, for the users with strong inter-beam interference, the adaptive frequency reuse can achieve similar data rate to fixed frequency reuse with $f = 1/2$ by efficiently eliminating the dominant inter-beam interference. For the users with low interference, the adaptive frequency reuse scheme can achieve similar performance to universal frequency reuse by using full frequency band for transmission.

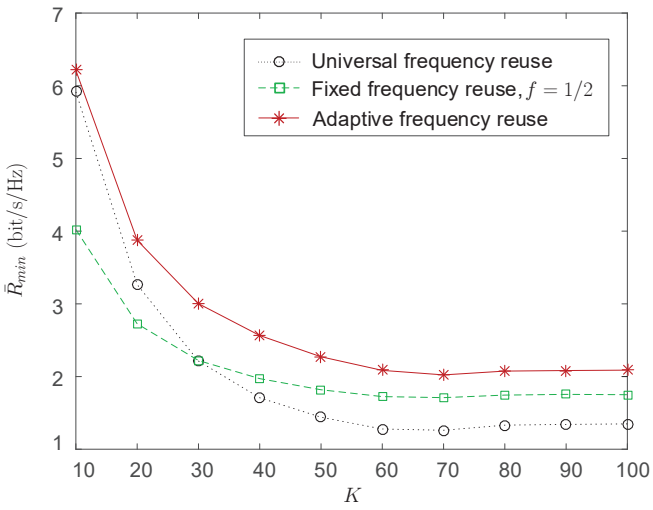
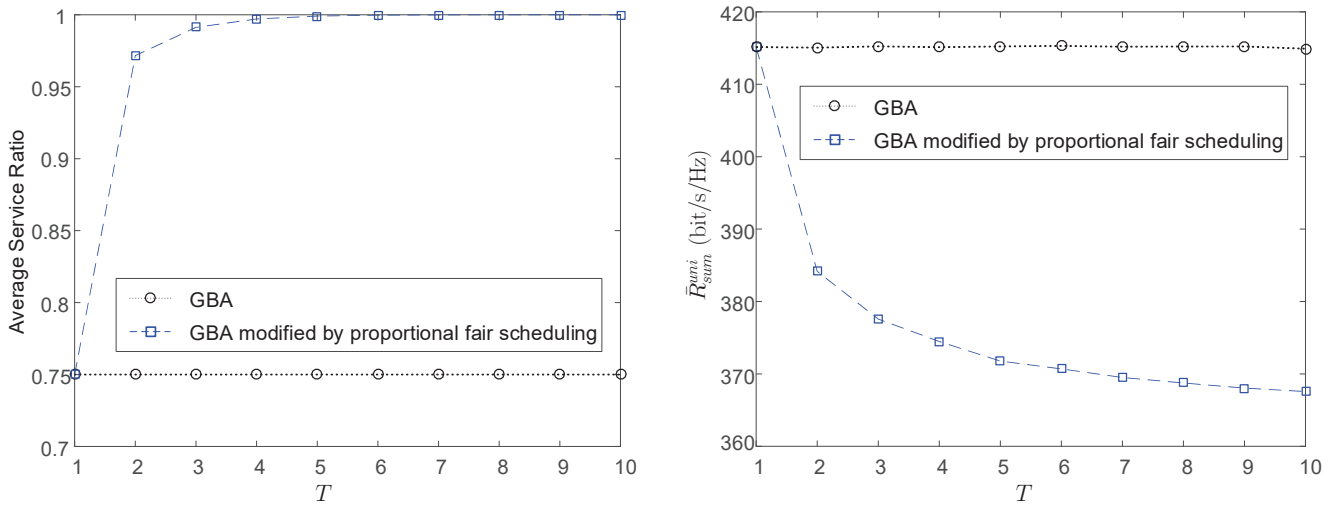


Fig. 15. Average minimum data rate of the served users $\bar{R}_{min}$ with the proposed adaptive frequency reuse, universal frequency reuse and fixed frequency reuse with $f = 1/2$. $\alpha = 2.2$, $\bar{P}_t/\sigma_0^2 = 20$dB, $N = 512$, $N_{RF} = 60$.

C. Simulation Results

Fig. 14 presents the average achievable data rate per worst-case user $\bar{R}_{worst} \triangleq \mathbb{E}_{\{\mathbf{r}_k:\ k\in\mathcal{K}\}}\left[R_k \mid k \in \mathcal{K}_{worst}\right]$ and the average sum data rate $\bar{R}_{sum} \triangleq \mathbb{E}_{\{\mathbf{r}_k:\ k\in\mathcal{K}\}}\left[\sum_{k\in\mathcal{K}_s} R_k\right]$ under varying number of users K. It can be seen in Fig. 14(a) that with universal frequency reuse, the average data rate per worst-case user $\bar{R}_{worst}^{uni}$ remains as a constant as the number of users K increases. The reason is that with universal frequency reuse, the strongest adjacent interfering beam of each worst-case user is allocated with the same frequency band, which contributes



(a) Average Service Ratio (b) $\bar{R}_{sum}^{uni}$

Fig. 16. (a) Average service ratio and (b) average sum data rate $\bar{R}_{sum}^{uni}$ with both the proposed original GBA algorithm and the modified GBA algorithm by considering proportional fair scheduling $\alpha = 2.2$, $\bar{P}_t/\sigma_0^2 = 20$dB, $N = 512$, $K = 80$, $N_{RF} = 60$.

to most of the worst-case user's inter-beam interference. As a result, by increasing the number of users K, although the total number of active beams increases, the inter-beam interference suffered by each worst-case user is almost a constant and thus the average data rate per worst-case user $\bar{R}_{worst}^{uni}$ is independent of the number of users K.

If fixed or adaptive frequency reuse is applied, the strongest interfering beam of a worst-case user is allocated with distinct half frequency band, resulting in no inter-beam interference for the worst-case user. That is, a worst-case user only suffers from inter-beam interference from other active beams. Since the number of served users/active beams K_s in the system first increases with the number of users K and then converges at the number of RF chains $K_s = N_{RF} = 60$, the inter-beam interference first increases and then converges. As a result, $\bar{R}_{worst}^{fix}$ and $\bar{R}_{worst}^{adp}$ first decrease and then become saturated. As can be expected from Fig. 13, it is shown in Fig. 14(a) that both fixed frequency reuse and adaptive frequency reuse can greatly improve the average data rate per worst-case user while fixed frequency reuse achieves a slightly higher data rate as some non-dominant interfering beams are allocated with distinct half frequency band which leads to lower inter-beam interference for the worst-case users. However, as these non-dominant interfering beams serve some non-worst-case users, a low average sum data rate with fixed frequency reuse factor $f = 1/2$ can be observed from Fig. 14(b).

Fig. 15 presents the average minimum data rate of the served users $\bar{R}_{min} \triangleq \mathbb{E}_{\{\mathbf{r}_k:\ k\in\mathcal{K}\}}\left[\min_{k\in\mathcal{K}_s} R_k\right]$. It can be clearly seen from this figure that the proposed adaptive frequency reuse achieves higher average minimum data rate than both fixed frequency reuse and universal frequency reuse, implying that max-min fairness can be improved by adopting the proposed adaptive frequency reuse scheme. In fact, although the worst-case users' data rates can be highly enlarged with fixed frequency reuse as Fig. 14(a) shows, other served users' data rates are also highly reduced due to the reduction of the communication bandwidth which might be even lower than the worst-case users' data rates, resulting in a lower average minimum data rate than that with adaptive frequency reuse.

V. Discussion

In this paper, the GBA algorithm has been proposed to maximize the sum data rate from the system's perspective in a single time slot, with the result that some users might not be served. In practical systems, however, the unserved users in a time slot will be given higher priority to be served in the following time slot to make sure that all the users in the system are successfully served eventually. It is therefore of practical importance to consider beam allocation and user scheduling over multiple time slots jointly. Specifically, our GBA algorithm can be modified by applying existing scheduling algorithms. For the sake of demonstration, a representative user scheduling algorithm, proportional fair scheduling [26], is adopted, which transmits to the user k^* with the largest $R_k(t)/\bar{R}_k(t-1)$ in each time slot t, where $R_k(t)$ is the data rate that user k can achieve in time slot t and $\bar{R}_k(t-1)$ is the average achieved data rate of user k till time slot $t-1$. To consider proportional fair scheduling, our original sum data rate maximization objective function of (8a) can be replaced by $\sum_{k\in\mathcal{K}} R_k^{uni}(t)/\bar{R}_k^{uni}(t-1)$ to maximize the weighted sum data rate, and our GBA algorithm can be modified accordingly.

Fig. 16 presents the average service ratio, i.e., the average ratio of the number of served users to the total number of users in the system over users' locations, and the average sum data rate of the system under varying number of considered contiguous time slots T. It is clearly shown in Fig. 16(a) that the average service ratio with the original GBA algorithm is lower than 1 as the number of RF chains, 60, is smaller than

the total number of users, 80, and thus maximum 60 users can be served, resulting in a service ratio of 0.75. By contrast, with the modified GBA algorithm by considering proportional fair scheduling the average service ratio increases first and then converges to 1 as number of considered time slots T increases, indicating that all the users in the system can be served as long as a sufficiently large number of time slots are considered for user scheduling. On the other hand, it is shown in Fig. 16(b) that the average sum data rate with proportional fair scheduling decreases with T and is smaller than that without proportional fair scheduling. The reason is that with proportional fair scheduling, higher priority is always given to the lower-data-rate users who might have bad channel conditions, resulting in a deteriorating average sum data rate with an increasing number of contiguous time slots T considered for user scheduling. We can then conclude from the above discussion that the user scheduling algorithm determines a tradeoff between sum data rate and fairness among users. How to design an optimal user scheduling scheme for the beam allocation based multiuser massive MIMO system is very interesting, which would be carefully studied in the future.

VI. Conclusion

In this paper, the beam allocation problem in a fixed-beam based multiuser massive MIMO system has been studied. By considering the practical constraint that the number of RF chains in the system is smaller than the number of users, a greedy beam allocation algorithm has been proposed to maximize the sum data rate with universal frequency reuse. However, simulation results reveal that although the proposed greedy algorithm can achieve near-optimal sum data rate, there could be a few worst-case users suffering from very severe inter-beam interference. Fixed frequency reuse has then been applied in the beam allocation based system to mitigate the strong inter-beam interference, which is shown to be able to improve the worst-case users' data rates at the cost of sacrificing the sum data rate. To harvest the benefits of universal frequency reuse and fixed frequency reuse, an adaptive frequency reuse scheme has been further proposed. Simulation results have shown that the proposed adaptive frequency reuse can increase the worst-case users' data rates while achieving similar sum data rate to universal frequency reuse. Additionally, it has been shown that by adopting adaptive frequency reuse the average minimum data rate of the served users can be improved, which indicates that the max-min fairness is enhanced among the served users.

The concept of adaptive frequency reuse proposed in this paper to mitigate inter-beam interference can be further explored to alleviate the inter-cell interference. How to design an efficient adaptive frequency reuse scheme to improve cell-edge users' data rates in a multi-cell massive MIMO system is an interesting topic, which deserves much attention in future work.

Appendix A
Definition of Uniform Matroids

Definition 2. *A uniform matroid $\mathcal{M} = (U, \mathcal{I})$ is a matroid on a finite ground set U with the independent set*

$$\mathcal{I} = \{X \subseteq U : |X| \leq m\}, \tag{A.1}$$

for some given parameter m.

Appendix B
Derivation of (15) and (16)

A. Derivation of (15)

Recall that the directivity of any beam $n \in \mathcal{B}$ with respect to an AoD θ is given in (1) by

$$D_n(\theta) = \frac{\sin^2\left(0.5N\pi\cos\theta - \beta_n\right)}{N\sin^2\left(0.5\pi\cos\theta - \frac{1}{N}\beta_n\right)}. \tag{B.1}$$

As illustrated in Fig. 6, the main direction of beam n, θ_n^m, is defined as the angle which maximizes the directivity $D_n(\theta)$, i.e.,

$$\theta_n^m = \arg\max_{\theta} D_n(\theta). \tag{B.2}$$

By combining (B.1) and (B.2), we have

$$0.5\pi\cos\theta_n^m - \frac{1}{N}\beta_n = 0. \tag{B.3}$$

It can be then obtained by substituting (2) into (B.3) that

$$\cos\theta_n^m = \frac{2n-1}{N} - 1. \tag{B.4}$$

B. Derivation of (16)

As Fig. 6 shows, the directivities of beam n and beam $n+1$ at the direction of their angular edge, θ_n^c, are equal. That is,

$$D_n(\theta_n^c) = D_{n+1}(\theta_n^c). \tag{B.5}$$

Since $\sin^2(x) = \sin^2(-x)$, it can be obtained from (B.1) and (B.5) that

$$0.5\pi\cos\theta_n^c - \frac{1}{N}\beta_n + 0.5\pi\cos\theta_n^c - \frac{1}{N}\beta_{n+1} = 0. \tag{B.6}$$

By substituting (2) into (B.6), we have

$$\cos\theta_n^c = \frac{2n}{N} - 1. \tag{B.7}$$

Appendix C
Derivation of (26) and (27)

Recall that the optimization problem formulated in (23a)–(23b) is given by

$$\max \quad \Delta\psi_{th} \tag{C.1a}$$

$$\text{subject to} \quad R_k^{adp} \geq R_k^{uni},\ \forall k \in \mathcal{K}_{worst}. \tag{C.1b}$$

For any user $k \in \mathcal{K}_s$, it is served by beam $n_k^{(1)}$, and its ith best beam with the ith largest directivity is denoted as beam $n_k^{(i)}$. The received desired signal power P_k of user k can be obtained according to (3) as

$$P_k = p_{n_k^{(1)}} \cdot D_{n_k^{(1)}}(\theta_k) \cdot \rho_k^{-\alpha}. \tag{C.2}$$

For any worst-case user $k \in \mathcal{K}_s$, as its second best beam $n_k^{(2)}$ is active, user k suffers from very strong inter-beam interference from beam $n_k^{(2)}$ with universal frequency reuse, which contributes to most of the interference I_k^{uni} received at user k. The inter-beam interference I_k^{uni} given in (7) can be then approximated as

$$I_k^{uni} \approx p_{n_k^{(2)}} \cdot D_{n_k^{(2)}}(\theta_k) \cdot \rho_k^{-\alpha}. \tag{C.3}$$

By combining (5), (6), (C.2) and (C.3), we have

$$\begin{aligned} R_k^{uni} &= \log_2\left(1 + \frac{\frac{P_t}{K_s} \cdot D_{n_k^{(1)}}(\theta_k) \cdot \rho_k^{-\alpha}}{\sigma_0^2 + \frac{P_t}{K_s} \cdot D_{n_k^{(2)}}(\theta_k) \cdot \rho_k^{-\alpha}}\right) \\ &\overset{P_t/N_0 \gg 1}{\approx} \log_2\left(1 + \frac{D_{n_k^{(1)}}(\theta_k)}{D_{n_k^{(2)}}(\theta_k)}\right). \end{aligned} \tag{C.4}$$

In contrast to the case with universal frequency reuse, each worst-case user k's serving beam $n_k^{(1)}$ and its strongest potential interfering beam $n_k^{(2)}$ are allocated with distinct frequency bands by applying the proposed adaptive frequency reuse scheme and thus user k is not interfered by beam $n_k^{(2)}$ any more. However, it still suffers from inter-beam interference from other active beams using the same half frequency band as beam $n_k^{(1)}$ for data transmission. To satisfy the constraint given in (C.1b), the potential minimum data rate achieved by user k, $\tilde{R}_k^{adp}$, should be larger than the data rate R_k^{uni}, i.e.,

$$\tilde{R}_k^{adp} \geq R_k^{uni}. \tag{C.5}$$

Note that the potential minimum data rate $\tilde{R}_k^{adp}$ with adaptive frequency reuse is achieved when beam $n_k^{(4)}$ is allocated with the same half frequency band as beam $n_k^{(1)}$.[4] Similar to (C.3), the corresponding inter-beam interference $\tilde{I}_k^{adp}$ can be obtained as

$$\tilde{I}_k^{adp} \approx p_{n_k^{(4)}} \cdot D_{n_k^{(4)}}(\theta_k) \cdot \rho_k^{-\alpha}. \tag{C.6}$$

By substituting (C.2) and (C.6) into (24), we have

$$\begin{aligned} \tilde{R}_k^{adp} &= \frac{1}{2}\log_2\left(1 + \frac{\frac{P_t}{K_s} \cdot D_{n_k^{(1)}}(\theta_k) \cdot \rho_k^{-\alpha}}{\frac{1}{2}\sigma_0^2 + \frac{P_t}{K_s} \cdot D_{n_k^{(4)}}(\theta_k) \cdot \rho_k^{-\alpha}}\right) \\ &\overset{P_t/N_0 \gg 1}{\approx} \frac{1}{2}\log_2\left(1 + \frac{D_{n_k^{(1)}}(\theta_k)}{D_{n_k^{(4)}}(\theta_k)}\right). \end{aligned} \tag{C.7}$$

Without loss of generality, let us assume

$$n_k^{(1)} = n, \tag{C.8}$$

and

$$n_k^{(2)} = n + 1, \tag{C.9}$$

i.e., $\psi_k \in [\psi_n^m, \psi_n^c]$. Then we have

$$\Delta\psi_k = \frac{2n}{N} - 1 - \psi_k, \tag{C.10}$$

[4]Although beam $n_k^{(3)}$ is the other adjacent beam of beam $n_k^{(1)}$, the inter-beam interference suffered by beam $n_k^{(1)}$ from beam $n_k^{(3)}$ is smaller than that from beam $n_k^{(4)}$ as either distinct half frequency band from that allocated to beam $n_k^{(1)}$ or full frequency band will be allocated to beam $n_k^{(3)}$ which results in no inter-beam interference or low inter-beam interference for user k.

according to (16)–(19), (C.8) and (C.9), and

$$n_k^{(4)} = n + 2. \tag{C.11}$$

By combining (1)–(2), (14), (C.4), (C.8) and (C.9), it can be obtained that

$$R_k^{uni} = \log_2\left(1 + \frac{\sin^2\left(\frac{\pi}{2}\psi_k - \frac{2n-N+1}{2N}\pi\right)}{\sin^2\left(\frac{\pi}{2}\psi_k - \frac{2n-N-1}{2N}\pi\right)}\right), \tag{C.12}$$

which can be further reduced to

$$\begin{aligned} R_k^{uni} &= \log_2\left(1 + \frac{\sin^2\left(-\frac{\pi}{2}\Delta\psi_k - \frac{\pi}{2N}\right)}{\sin^2\left(-\frac{\pi}{2}\Delta\psi_k + \frac{\pi}{2N}\right)}\right) \\ &\overset{\text{for large } N}{\approx} \log_2\left(1 + \frac{\left(-\frac{\pi}{2}\Delta\psi_k - \frac{\pi}{2N}\right)^2}{\left(-\frac{\pi}{2}\Delta\psi_k + \frac{\pi}{2N}\right)^2}\right), \end{aligned} \tag{C.13}$$

according to (C.10). Similarly, we have

$$\begin{aligned} \tilde{R}_k^{adp} &= \frac{1}{2}\log_2\left(1 + \frac{\sin^2\left(-\frac{\pi}{2}\Delta\psi_k - \frac{3\pi}{2N}\right)}{\sin^2\left(-\frac{\pi}{2}\Delta\psi_k + \frac{\pi}{2N}\right)}\right) \\ &\overset{\text{for large } N}{\approx} \frac{1}{2}\log_2\left(1 + \frac{\left(-\frac{\pi}{2}\Delta\psi_k - \frac{3\pi}{2N}\right)^2}{\left(-\frac{\pi}{2}\Delta\psi_k + \frac{\pi}{2N}\right)^2}\right). \end{aligned} \tag{C.14}$$

By substituting (C.13) and (C.14) into (C.5), it can be obtained that

$$\Delta\psi_k \leq \left(\sqrt{\frac{2}{C} - C^2 - 1} - C\right) \cdot \frac{1}{N}, \tag{C.15}$$

with

$$C = \sqrt{\frac{1}{3}\left(\left(17 + 3\sqrt{33}\right)^{\frac{1}{3}} - 2\left(17 + 3\sqrt{33}\right)^{-\frac{1}{3}} - 1\right)}. \tag{C.16}$$

As $\Delta\psi_k \leq \Delta\psi_{th}$ for any worst-case user $k \in \mathcal{K}_{worst}$ according to Definition 1, we have

$$\Delta\psi_{th} \leq \left(\sqrt{\frac{2}{C} - C^2 - 1} - C\right) \cdot \frac{1}{N}. \tag{C.17}$$

Finally, (26) and (27) can be obtained by combining (C.1a), (C.16), and (C.17).

ACKNOWLEDGMENT

This work is currently supported by the UK Engineering and Physical Sciences Research Council, project EP/L026031/1, NIRVANA. It was also partly carried out within the framework of the EU-Japan Horizon 2020 Research and Innovation Programme, grant agreement no. 643297 (RAPID).

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[5] Y. Xin, D. Wang, J. Li, H. Zhu, J. Wang, and X. You, "Area spectral efficiency and area energy efficiency of massive MIMO cellular systems," *IEEE Trans. Veh. Tech.*, vol. 65, no. 5, pp. 3243–3254, May 2016.

[6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[7] D. Wang, Y. Zhang, H. Wei, X. You, X. Gao, and J. Wang, "An overview of transmission theory and techniques of large-scale antenna systems for 5G wireless communications," *Sci. China Inf. Sci.*, vol. 59, p. 081301, pp. 1–18, Aug. 2016.

[8] H. Wei, D. Wang, H. Zhu, J. Wang, S. Sun, and X. You, "Mutual coupling calibration for multiuser massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 606–619, Jan. 2016.

[9] V. Venkateswaran and A. Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.

[10] O. N. Alrabadi, E. Tsakalaki, H. Huang, and G. F. Pedersen, "Beamforming via large and dense antenna arrays above a clutter," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 314–325, Feb. 2013.

[11] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[12] J. Wang, H. Zhu, L. Dai, N. J. Gomes, and J. Wang, "Low-complexity beam allocation for switched-beam based multiuser massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8236–8248, Dec. 2016.

[13] J. Brady and A. Sayeed, "Beamspace MU-MIMO for high-density gigabit small cell access at millimeter-wave frequencies," in *Proc. IEEE SPAWC*, pp. 80–84, June 2014.

[14] P. Amadori and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Trans. Wireless Commun.*, vol. 63, no. 6, pp. 2212–2223, June 2015.

[15] J. Hogan and A. Sayeed, "Beam selection for performance-complexity optimization in high-dimensional MIMO systems," in *Proc. CISS*, pp. 337–342, Mar. 2016.

[16] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmWave massive MIMO systems," *IEEE Commun. Letters,* vol. 20, no. 5, pp. 1054–1057, May 2016.

[17] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[18] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Select. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[19] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Letters*, vol. 3, no. 6, pp. 653–656, Dec. 2014.

[20] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems," in *Proc. IEEE ICASSP*, pp. 2929–2933, Apr. 2015.

[21] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[22] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1952–1967, May 2016.

[23] J. Butler and R. Lowe, "Beam-forming matrix simplifies design of electrically scanned antennas," *Electronic Design,* Apr. 1962.

[24] R. C. Hansen, *Phased Array Antennas*, John Wiley & Sons, Inc., 2009.

[25] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[26] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

Junyuan Wang (S'13-M'15) received the B.S. degree in Telecommunications Engineering from Xidian University, Xi'an, China, in 2010, and Ph.D. degree in Electronic Engineering from City University of Hong Kong, Hong Kong, China, in 2015.

She is currently a Research Associate at the School of Engineering and Digital Arts, University of Kent, Canterbury, U.K. Her research interests include performance analysis, massive MIMO and distributed antenna systems/C-RAN.

Huiling Zhu (M'04) received the B.S. degree from Xidian Univeristy, Xian, China, and the Ph.D. degree from Tsinghua University, Beijing, China. She is currently a Reader (Associate Professor) with the School of Engineering and Digital Arts, University of Kent, Canterbury, U.K. Her research interests are in the area of broadband wireless mobile communications, covering topics, such as radio resource management, distributed antenna systems, MIMO, cooperative communications, device-to-device communications, small cells, and heterogeneous networks. She was a recipient of the Best Paper Award from the IEEE Globecom 2011, Houston. She has participated in a number of European and industrial projects in these topics. She was holding European Commission Marie Curie Fellowship from 2014 to 2016. She has served as the Publication Chair for the IEEE WCNC 2013, Shanghai, Operation Chair for the IEEE ICC 2015, London, and Symposium Co-Chair for the IEEE Globecom 2015, San Diego. She currently serves as an Editor for the IEEE Transactions on Vehicle Technology.

Nathan J. Gomes (M92SM06) received the B.Sc. degree in electronic engineering from the University of Sussex, Sussex, U.K., in 1984, and the Ph.D. degree in electronic engineering from University College London, London, U.K., in 1988.

From 1988 to 1989, he held a Royal Society European Exchange Fellowship with ENST, Paris, France. Since late 1989, he has been with the University of Kent, Canterbury, U.K., where he is currently a Professor of Optical Fibre Communications. His current research interests include fiber-wireless access, and the fronthaul for future mobile networks.

He was the TPC Chair for IEEE International Conference on Communications, London, U.K., 2015.

Jiangzhou Wang (M'91-SM'94-F'17) is currently a Professor at the University of Kent, U.K. He has authored over 300 papers in international journals and conferences in the areas of wireless mobile communications and three books. He is an IEEE Fellow and IET Fellow. He received the Best Paper Award from IEEE GLOBECOM2012 and was an IEEE Distinguished Lecturer from 2013 to 2014. He is the Technical Program Chair of IEEE ICC2019 in Shanghai. He was the Executive Chair of IEEE ICC2015 in London and the Technical Program Chair of IEEE WCNC2013. He was an Editor for IEEE Transactions on Communications from 1998 to 2013 and was a Guest Editor for IEEE Journal on Selected Areas in Communications, IEEE Communications Magazine, and IEEE Wireless Communications. Currently, he is an editor of Science China Information Sciences. His research interests include massive MIMO, Cloud RAN, NOMA, and D2D communications.